Longitudinal genetic analysis of problem behaviours in biologically related and unrelated adoptees

van der Valk, J.C.; Verhulst, F.C.; Neale, M.C.; Boomsma, D.I.

published in
Behavior Genetics
1998

DOI (link to publisher)
10.1023/A:1021621719059

document version
Publisher's PDF, also known as Version of record

Behavior Genetics, Vol. 28, No. 5, 1998

Longitudinal Genetic Analysis of Problem Behaviors in Biologically Related and Unrelated Adoptees

Jolande C. van der Valk,[1,3,4] Frank C. Verhulst,[1] Michael C. Neale,[2] and Dorret I. Boomsma[3]

Received 9 April 1997—Final 3 August 1998

The genetic and environmental influences on problem behaviors at two assessment points, three years apart, and their stability were studied in a sample of international adoptees, initially aged 10 to 15 years. Parents of 111 pairs of adopted biological siblings, 221 pairs of adopted nonbiological siblings and 1484 adopted singletons completed the Child Behavior Checklist (75 pairs, 154 pairs and 1080 singletons respectively at second assessment). At first assessment, genetic factors accounted for more than 50% of the variance in the Externalizing, Aggressive Behavior, Attention Problems and Social Problems scales. Shared environmental influences explained 40% of the variance in the Total Problem scale and less for all other scales. Nonshared environmental influences were most important for the Internalizing scale and its subscales, and for the Thought Problems and Delinquent Behavior scales. At the second assessment, genetic factors explained most of the variance in the Total Problem, Externalizing and Aggressive Behavior scales, while nonshared environmental influences explained most of the variance in all other scales. Shared environmental influences explained 33% of the variance in the Internalizing scale and less for the other scales. The stability of the Externalizing scale over time was caused mostly by genetic factors, while nonshared environmental factors mostly caused the stability of the Internalizing scale.

KEY WORDS: Problem behaviors; longitudinal analysis; adolescent psychopathology; Child Behavior Checklist; international adoptees; behavior genetics.

INTRODUCTION

Prospective studies showed high stability of behavioral and emotional problems during childhood, adolescence and early adulthood. Across studies, one-third to one-half of children with initial deviant scores maintain deviant scores across 2- to 6-year intervals (Koot, 1995). The presence of multiple problems increases the likelihood of stability. Age and gender of the child do not seem to be of major influence. Although most children show fluctuations over time in the level of deviant behavior, extreme changes are rare. There are indications that, at least from school age onwards, the stability of problem behavior is specific. Higher stabilities over time are reported within than across problem areas. Using the Child Behavior Checklist (CBCL) (Achenbach, 1991a,b), Externalizing scores were more predictive of later Externalizing scores than of Internalizing scores, and vice versa (Verhulst and van der Ende, 1993). When using rating scales,

[1] Department of Child and Adolescent Psychiatry, Sophia Children's Hospital, Rotterdam, The Netherlands.
[2] Department of Psychiatry, Medical College of Virginia/Virginia Commonwealth University, Richmond, Virginia 23298, U.S.A.
[3] Department of Psychonomics, Free University, Amsterdam, The Netherlands.
[4] To whom correspondence should be addressed at the department of Child and Adolescent Psychiatry, Sophia Children's Hospital, Rotterdam, The Netherlands. Telephone: (31) 10-4636092. Fax: (31) 10-4636803. E-mail: vandervalk @psys.azr.nl.

0001-8244/98/0900-0365$15.00/0

internalizing problems generally are almost as stable as externalizing problems. However, when clinical diagnoses are made, emotional disorders seem to show better prognosis than conduct or hyperactivity disorders. That is, children who persist in their deviant behavior tend to show oppositional, aggressive or antisocial behaviors, whereas the majority of children who improve initially showed fearful, inhibited, or depressive behaviors (Esser *et al.*, 1990; Verhulst *et al.*, 1993).

Given the stability of problem behaviors, the next question is what the etiology of this stability is. Problem behaviors of children generally involve quantitative variations in behavior that most children display to some degree. These continuous variations in behavioral problems are hypothesized to be caused by multiple genes and environmental influences. The effects of genes and environment on variation in behavior can be studied with genetically informative subjects such as twins or adoptees. Likewise, the contributions of genetic and environmental factors to the covariation of behavior across time can be assessed with genetically informative subjects who are measured repeatedly across time. We have studied the etiology of problem behaviors during adolescence over a 3 year interval in a sample of internationally adopted children. In this sample of biologically related and unrelated adopted siblings and singletons, the stability and change of genetic and environmental influences on different problem behaviors were assessed, using the Child Behavior Checklist (CBCL/4-18, Achenbach, 1991a) to obtain parental ratings of behavioral problems.

Longitudinal studies can resolve whether changes in heritability during adolescence are due to changes in genetic or environmental variances with age. More importantly, however, longitudinal studies can reveal how genes and environmental influences operate throughout development. For example, is an increase in heritability due to new, additional, genetic factors being expressed as children grow older, or is there an amplification of existing genetic influences? This second objective addresses the question to what extent phenotypic stability is due to the same genes being expressed at different ages and to what extent phenotypic stability is due to the same environmental influences being of importance. Contrary to popular points of view, genetically determined characters need not be stable, nor are longitudinally stable characters always influenced by heredity (Molenaar, Boomsma & Dolan, 1991).

Several studies have discussed the importance of genetic and environmental influences on children's problem behaviors (see Edelbrock *et al.*, 1995). However, we know of only one study that has prospectively assessed the stability and change of genetic and environmental influences on children's problem behaviors. Schmitz *et al.* (1995) collected CBCL data over a 5-year period for children who were almost 3 years old at the first assessment. In their relatively small longitudinal sample of 95 twin pairs, Schmitz *et al.* (1995) found that the same genes were operating at both the earlier and the later time point for the Aggressive Behavior scale. For the Somatic Complaints and Anxious/Depressed scales some genetic influences persisted, but newly expressed genetic variation during middle childhood had a greater impact. Shared environmental influences remained the same for all CBCL scales in early and middle childhood, although these influences only explained a significant proportion of the observed variances of the Internalizing scale and the Total Problem scale. As Schmitz *et al.* (1995) indicated, these interesting results should be replicated with a larger sample before definite conclusions can be drawn.

The present study comprises three groups: a group of 111 pairs of biologically related siblings, adopted into the same family; a group of 221 pairs of nonbiologically related siblings, also adopted into the same family; and a group of 1484 singly adopted adolescents. At the second assessment 75 pairs, 154 pairs and 1080 singletons, respectively, participated again. Adoptees were aged 10 to 15 years at the first assessment; 95.9% of the sample was between 11 and 14 years of age. When they were assessed again three years later using the same instrument, 95.8% of the sample was between 14 and 17 years of age. A special feature of our study is that all groups are raised by adoptive parents. In most other studies adopted children are compared with controls who are raised by their biological parents. Of course, having the status of 'adopted child' or not can have a profound influence on the measured variables. For some adopted children, this status might be difficult to accept, thereby possibly increasing the amount of problem behaviors shown. In this study, using solely adopted children, results can not be distorted by this interference. Also, biological children can show different

estimates of genetic and environmental influences than adopted children do, because of possible interactions between the genotype and environment of parents and their children. Using only adopted children, we will be able to measure genetic and environmental estimates that are not distorted by this type of genotype-environment interaction.

Van den Oord *et al.* (1994) used the same adoption sample to determine the heritability of different problem behaviors at the first assessment. The authors did not use the scales as constructed by Achenbach (1991a,b), but developed their own scales that differed slightly from those by Achenbach. Van den Oord *et al.* (1994) found that the Internalizing scale showed almost no genetic influences. Nonshared environmental influences accounted for almost all of the variance. However, the Externalizing scale showed genetic effects that were larger than either nonshared or shared environmental influences. Van den Oord *et al.* (1994) found that variation in behavioral problems was neither influenced by the number of siblings, nor by the influence of siblings interacting with each other. Sex differences in heritability were found for most problem behaviors, showing larger genetic influences for boys (but the effect was small).

The goals of the current study were first, to estimate at two assessment points during adolescence the genetic and environmental influences on different problem behaviors. Second, to examine the continuity and change of these influences over the 3-year interval, addressing the question to what extent the genetic and environmental factors, expressed at the first assessment, remain important over time and to what extent new genetic and environmental factors become of importance.

METHODS

Assessment Instrument

The CBCL (Achenbach, 1991a,b) consists of 20 competence items and 120 problem items. Only the problem items were used in this study. They were scored by the parents on a 3–point scale based on the occurrence of the behavior during the preceding 6 months: 0 if the problem item was not true of the child, 1 if the item was somewhat or sometimes true, and 2 if it was very true or often true. Using factor analyses, Achenbach (1991a) computed eight syndrome scales from these 120 problem items. The syndrome scales were named: Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Delinquent Behavior and Aggressive Behavior. The first three syndrome scales were summed to form a broad-band grouping, called Internalizing. The last two syndrome scales were summed to form a broad-band grouping called Externalizing. The Total Problem scale was computed by summing the scores given to the 120 problem items, with the exception of 2 problem items concerning allergy and asthma.

The good reliability and validity of the CBCL (Achenbach, 1991b) was confirmed for the Dutch version of the CBCL (Verhulst *et al.*, 1985; 1996). The test-retest reliability over a period of 2 weeks, measured in 89 children chosen at random from the Dutch population, was highest for the Total Problem scale (Pearson correlation of 0.91) and lowest for the Thought Problems scale (0.74), with correlations significant ($p < 0.001$) (Verhulst *et al.*, 1996). De Groot *et al.* (1994) studied the cross-cultural generalizability of the Dutch version of the CBCL. Confirmatory factor analysis of the American syndromes in a sample of 2335 clinically referred Dutch children, aged 4 to 18 years, strongly supported the generalizability of the CBCL. In a sample consisting of 4- to 16-year-olds, drawn in 1983 from the Dutch province of Zuid-Holland (see Verhulst, *et al*, 1985), the stability over a four year interval was highest for the Aggressive Behavior scale (Pearson correlation of 0.65) and lowest for the Thought Problems scale (0.24). The Total Problem scale showed a four year stability of 0.64. Over a two year interval, the Pearson correlations for most scales were higher (Verhulst *et al., 1996*).

The distribution of the summed scores on the different scales was skewed, because most adoptees showed either none or just a few behavior problems. Logarithmic transformations were applied to reduce skewness. After transformation, only the Somatic Problems and Thought Problems scales showed a skewness larger than 1.0 and only the Thought Problems scale showed a kurtosis larger than 1.3. These were the only scales deviating from normality, and they did so at both assessments.

Subjects

The prevalence of problem behaviors in adoptees was assessed twice, with a mean interval of

3.2 years (SD of 2.5 months). The original sample at the first assessment was selected from the central adoption register of the Dutch Ministry of Justice in 1986. It consisted of 3519 children, legally adopted by nonrelatives in the Netherlands and born outside the Netherlands between January 1, 1972 and December 31, 1975. Dutch adoption agency policies do not include selective placement. The adoptive parents were asked by letter to participate in the study. If they consented, a prepaid return envelope, a CBCL with instructions and a questionnaire about the history and health of the child were sent. If any help was needed, the parents were instructed to phone the investigators.

From the original sample, 162 adoptees had moved abroad, 39 were untraceable and 9 had died. Of the 3309 adoptees whose parents were sent the questionnaires, 2148 (64.9%) usable CBCLs were returned by mail; parents of 238 adoptees refused to participate and on 923 adoptees no response was received. For reasons of privacy, it was not permitted to contact the nonresponders or collect relevant data on them from the original adoption files. For this reason, responders, explicit refusers and nonresponders could only be compared on the adoptee's sex, actual age and age at placement in the adoptive home. The only difference found was a tendency of parents of adoptees that were placed in the adoption home at relatively later ages, to respond more than parents of adoptees that were placed in the adoption home at relatively earlier ages (Verhulst *et al.*, 1990). The respondents consisted of 45.4% mothers, 23% fathers, 28.5% mothers and fathers together filling in one questionnaire and 3.1% others (like the adoptee him/herself filling in, or assisting the parents with filling in, the questionnaire). Parental occupation was measured on a 6-step scale (van Westerlaak *et al.*, 1975). When both parents were employed, the highest level of one of them was used. The distribution of parental occupation was: 9.1% low (occupational levels 1 and 2); 25.8% middle (levels 3 and 4) and 65.1% high (levels 5 and 6). The majority of adoptive parents had a higher level of occupation (mean of 4.61, SD of 1.40). The distribution of adoptees across native countries was: Korea 32.0%, Colombia 14.6%, India 9.5%, Indonesia 7.9%, Bangladesh 6.7%, Lebanon 4.9%, Austria 5.0%, other European countries 4.2%, other non-European countries 15.2%. For the current study, the responders were divided into three groups: one group of 222 adolescents who were biologically related and adopted together into the same home, one group of 442 adolescents who were not biologically related but also adopted together into the same home and one group of 1484 adolescents who were adopted singly. There was never more than one pair of siblings in a family.

The responders of the first assessment were contacted again three years later. Of this group, 29 adoptees were untraceable, 8 had moved abroad, 3 had died and 37 were not approached because they were participating in another study. Parents of 2071 adoptees were sent the CBCL and a questionnaire about the general functioning of their adopted child(ren). A reminder was sent to the nonresponders and those who still did not respond were telephoned. Usable CBCLs were received from 1538 adoptees (74%). Adoptive parents that did not respond had adoptees that were slightly older and had slightly higher problem scores at the first assessment (Verhulst and Versluis-den Bieman, 1995). The respondents consisted of 61.3% mothers, 19.8% fathers and 18.9% mothers and fathers together filling in one questionnaire. The category others did not occur at the second assessment. Parental occupation and the distribution across native countries had not changed. For the current study, the responders were divided again into three groups: one group of 150 biologically related siblings, one group of 308 nonbiologically related siblings and one group of 1080 singly adopted adolescents.

Table I shows the relation between country of origin and the groups of biological siblings, nonbiological siblings and singletons. The chi-square test was significant at both assessments, indicating that there were differences among the three groups concerning their countries of origin. More biologically related siblings came from Korea and Columbia versus other Asian countries and Europe. Within pairs, biologically related siblings and 78% of the nonbiologically related siblings came from the same country of origin. Within the group of nonbiological adoptees, siblings who came from different countries tended to be somewhat more physically neglected before their placement in the adoptive home than siblings who had the same country of origin. For all other measured characteristics no differences were found.

Table II shows the tests for differences between the groups of biological siblings, nonbiol-

Table I. Countries of Origin of Adoptees in Percentages[a]

	Biological		Non-biological		Singletons	
Country of origin	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Korea	47.3	53.3	21.5	22.7	32.9	34.4
Colombia	26.6	22.0	13.3	14.0	13.2	13.3
India	2.7	2.7	10.0	11.7	10.3	10.5
Indonesia	9.9	8.7	6.6	6.8	8.0	7.8
Bangladesh	2.7	2.7	4.1	3.6	8.0	8.2
Lebanon	0.0	0.0	12.2	13.0	3.4	3.5
Austria	2.7	4.0	9.7	9.7	3.9	3.8
Other European	0.0	0.0	4.5	4.2	4.8	3.5
Other non-European	8.1	6.7	18.1	14.3	15.4	15.0
Number of children	222	150	442	308	1484	1080

[a] The chi-square (χ^2) test showed significant differences between the groups at both assessment points (Time 1: $\chi^2 = 202.08$, df = 16, p = .000; Time 2: $\chi^2 = 140.59$, df = 16, p = .000).

ogical siblings and singletons for different background characteristics at both assessments. The chi-square test showed a significant difference between the groups at both assessment points in number of changes in the caretaking environment that the child experienced before it was adopted. Biological siblings had experienced more changes than singletons, who had experienced more changes than nonbiological siblings. No significant differences were found between the three groups in whether the child had been physically neglected or abused before placement in the adoptive home. Comparing the groups by age of placement in the adoptive home, the chi-square test did show a significant difference at both assessment points. Biological siblings were placed in their adoptive homes later than singletons, who were placed later than nonbiological siblings. The child's physical health at the time of placement did not show a significant difference between the groups. ANOVA was used as a test of group differences for the age of the adoptee at measurement. Although shown in the table in years, we used a measurement in months. Only at the second assessment did ANOVA show a difference between the groups in age. Post-hoc tests indicated that the biological siblings were significantly older than the singletons. Parental occupation, of which scale 1 and 2 of van Westerlaak *et al.* (1975) were taken together (constituting the category 'low' occupation) to get the right expected frequencies in the cells, did not show a difference between the groups.

Missing Data and Different Kinds of Raters

At the second assessment a response bias was found. Parents who cooperated again had indicated fewer problems for their adopted children at the first assessment than parents who did not cooperate again. For the nonbiologically related and the singly adopted adolescents, this response bias was found only for the Delinquent Behavior scale (means of 2.54 versus 3.26 and 2.98 versus 3.49, respectively). However, for the biologically related siblings, ANOVA showed the same significant difference for 6 of the 11 CBCL syndrome scales. The mean reported problems at the first assessment of second time responders versus dropouts were accordingly, Anxious/Depressed: 3.77 versus 4.80, Social Problems: 2.31 versus 3.30, Delinquent Behavior: 2.45 versus 3.66, Aggressive Behavior: 5.47 versus 7.24, Externalizing: 6.29 versus 8.06, Total Problem: 10.83 versus 12.54.

This missing-data pattern at the second assessment is related to variables that have been measured (the CBCL scales at assessment 1) and thus the pattern is included in the analysis (Graham *et al.*, 1997). Although the data is not "missing completely at random", the missing-data pattern is consistent with "missing at random" (Little and Rubin, 1989; Graham *et al.*, 1996). Whether the data are missing on the later assessments is, at least in part, predicted by variables that are not missing. An appropriate statistical technique to handle such data is based on the maximization of the likelihood of the observed data. The likelihood gives an indication of how good the theoretical model, with its estimated parameters, represents the observed data. Even when the data are not strictly missing at random, maximum likelihood often reduces nonresponse biases (Little and Rubin, 1989; Muthén *et al.*, 1987). We used Mx (Neale, 1997), a structural equation modeling program that allows estimation of the raw maximum likelihood function at the level of the individual (Graham *et al.*, 1997; Neale, 1997; Wothke and Arbuckle, 1995). This fitting function corrects for the nonresponse bias at the second assessment by calculating the appropriate mean vector and covariance matrix for each observation separately, using per observation all information available. By using raw maximum

Table II. The Means and Standard Deviations per Group, and the Test of Differences Between the Groups, for Three Conditions: Before Placement in the Adoptive Home, at Placement in the Adoptive Home, and at Time of Measurement, for Assessment 1 and 2

	Biological	Nonbiological	Singletons	F/χ^2	df	p
Assessment 1:						
Conditions before placement						
changes in caretaking	1.86 (0.62)	1.60 (0.63)	1.70 (0.62)	31.41	4	.000[b]
physical neglect	1.64 (0.76)	1.56 (0.71)	1.64 (0.75)	4.53	4	.339
physical abuse	1.20 (0.48)	1.14 (0.42)	1.17 (0.45)	3.09	4	.543
Conditions at placement						
age at placement	4.95 (1.84)	2.89 (1.92)	3.73 (2.21)	174.44	14	.000[b]
physical health	1.40 (0.49)	1.41 (0.49)	1.44 (0.50)	2.72	2, 2145	.257
Conditions at measurement						
age of adoptee[a]	12.5 (1.18)	12.4 (1.15)	12.3 (1.17)	2.52	2	.081
occupation of parents	4.71 (1.39)	4.63 (1.42)	4.59 (1.39)	11.28	8	.186
Assessment 2:						
Conditions before placement						
changes in caretaking	1.84 (0.58)	1.60 (0.61)	1.69 (0.59)	19.27	4	.001[b]
physical neglect	1.59 (0.76)	1.57 (0.72)	1.64 (0.75)	3.59	4	.464
physical abuse	1.25 (0.52)	1.13 (0.40)	1.17 (0.46)	7.37	4	.118
Conditions at placement						
age at placement	4.97 (1.79)	2.86 (1.82)	3.69 (2.20)	136.05	14	.000[b]
physical health	1.37 (0.49)	1.44 (0.50)	1.45 (0.50)	2.97	2	.227
Conditions at measurement						
age of adoptee	15.8 (1.17)	15.6 (1.20)	15.4 (1.16)	6.53	2, 1535	.002[b]
occupation of parents	4.77 (1.30)	4.67 (1.42)	4.63 (1.39)	8.21	8	.413

[a] The variable "age of adoptee", given in years, used a measurement in months.
[b] F or χ^2 test showed a significant difference between the groups.

likelihood, the likelihood of the theoretical model was calculated separately for each pedigree and subsequently maximized over the different pedigrees (Neale, 1997). Although the singletons' data did not give any information about the genetic or environmental influences, their data did provide information on the variances at the first and second assessment and on the covariance between the assessments.

Different raters might vary in their tendency of reporting problem behaviors; for instance, mothers might report more problems than fathers. In order to be able to correct for this rater bias, while having only one questionnaire (one kind of rater) per child, we allowed different kinds of raters to have different means for reported problem behaviors. This can be done in Mx by a feature called definition variables, allowing 'multilevel' statistical analyses. Mx extracts the definition variable, in this case 'kind of rater', from the data and restricts modeling, separately for each kind of rater, to the other variables (the CBCL scales). The usual raw data log-likelihood function is computed for the theoretical model, while using the appropriate mean matrix for each 'kind of rater'.

Model

A genetic model was fitted to the variances and covariances between siblings. Nonbiologically related siblings, who only resemble each other because of similar shared environmental influences, were compared with biologically related siblings, who can also resemble each other because they share on average half of their genes. By comparing the similarity between the biologically related adoptees with the similarity between the nonbiologically related adoptees, identification of the model to estimate the contributions of genotype (A), shared environment (C) and nonshared environment (E) is achieved. If the biologically related adoptees resemble each other to the same degree as the nonbiologically related adoptees, only environmental factors can be of importance in explaining sibling resemblance. However, when the biologically related adoptees resemble each other

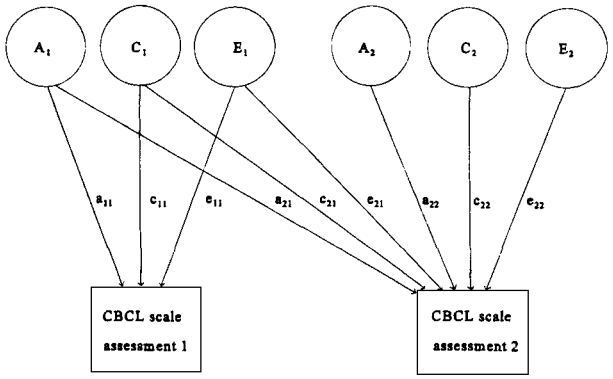

Fig. 1. Path diagram depicting the bivariate longitudinal ACE model. Latent variables A_1, C_1 and E_1 refer to the genetic, shared environmental and nonshared environmental factors that have an influence on the observed variance of the parental ratings at both assessment points. A_2, C_2, E_2 and correspond to the genetic, shared environmental and nonshared environmental factors that only have an influence at the second assessment point. The strength of the relative influences of the latent factors on the parental ratings are indicated by the path coefficients a_{11}, c_{11} and e_{11} for ratings at the first assessment and by a_{22}, c_{22} and e_{22} for ratings at the second assessment. The path coefficients a_{21}, c_{21} and e_{21} refer to the strength of the relative influences of the latent factors that retain their influences over time. The covariance between the two assessment points, for instance the genetic covariance, can be estimated by multiplying $a_{11}a_{21}$. The total covariance equals $a_{11}a_{21} + c_{11}c_{21} + e_{11}e_{21}$.

more than the nonbiologically related adoptees, genetic factors are supposed to be of importance, since the only difference between the two groups is in genetic relatedness.

To estimate the longitudinal genetic and environmental factors on the different CBCL scales, a bivariate Cholesky decomposition (Neale and Cardon, 1992) was fitted to the log-10 transformed, raw data. This model, shown in Fig. 1, decomposes the observed variance of the parental ratings into three latent factors that have, sequentially over time, an influence at both assessment points, i.e. genetic (A_1), shared environmental (C_1) and nonshared environmental (E_1) factors, as well as three latent factors that only have an influence at the second assessment point, i.e. genetic (A_2), shared environmental (C_2) and nonshared environmental (E_2) factors. The relative influences of the latent genetic and environmental factors on the different CBCL scales are indicated by the paths (i.e., a, c, e). To estimate the proportion of genetic (or environmental) influences that are active at the second assessment, the squared path of the first genetic (or environmental) factor has to be summed with the squared path of the second genetic (or environmental) factor and divided by the summed squared paths of the total amount of genetic and environmental influences at the second assessment. The genetic (or environmental) influences that are expressed only at the second assessment can be estimated by dividing the squared path from the second genetic (or environmental) factor by the total variance at the second assessment.

Using this bivariate Cholesky decomposition, it is possible to partition the covariance between the two assessment points into genetic and shared and nonshared environmental covariance. This indicates to which extent the stability of problem behaviors is caused by genes, shared, or nonshared environment. The percentage of covariance can be estimated by multiplying the path of the first latent factor, leading towards the scales measured at the first assessment, with the path leading from the first latent factor towards the scales measured at the second assessment and dividing this with the total covariance between the two assessment points.

Model fitting

An indication of how well a particular model fits the data is given by the likelihood. A good model is one that represents the observed results with a high likelihood. To be able to test the goodness of fit of the theoretical ACE model, the model's log-likelihood (LL) is subtracted from the LL of a less constrained model. By multiplying the result by 2, a chi-square test statistic is obtained. This chi-square test statistic indicates whether the theoretical ACE model describes the observed data adequately. The degrees of freedom for this test statistic are the number of parameters in the theoretical ACE model, subtracted from the number of parameters in the less constrained model.

First we tested whether the bivariate Cholesky model fitted the observed data significantly worse than a saturated model, which is a model without any constraints. The saturated model estimates the means and the variance-covariance matrices separately for each rater and for each group of adoptees. Second, when the bivariate Cholesky model did not show a significantly worse fit to the observed data, we tested whether the means between the three groups (biologically related adoptees, nonbiologi-

cally related adoptees and singletons) and between siblings (within the same family) could be constrained to be equal. Third we tested whether the means of the different kinds of raters and of the first and second assessment could be constrained to be equal. For each CBCL problem scale, the most simplified model for means was retained to analyze the causes of variation in the observed data and to test whether this best fitting ACE model could further be simplified by removing the genetic and shared and nonshared environmental factors. Finally, to get more detailed information about the precision of the genetic, shared environmental and nonshared environmental estimates, 95% likelihood-based confidence intervals were estimated. These confidence intervals are estimated by finding the maximum and minimum values of the path estimate that—with all other paths of the model still free to vary—cause a loss of fit equal to a chi-squared of 3.84 (Neale and Miller, 1997).

RESULTS

Description of the Data

Table III shows the means and standard deviations for singletons, oldest (first line) and youngest (second line) siblings and the correlations between the siblings for each group, at both assessments for all CBCL scales. The longitudinal correlations between the measurements are given in the last three columns. For all CBCL scales, the adoption sample showed higher means and higher standard deviations than the Dutch normative sample (Verhulst *et al.*, 1996). Within the adoption sample, the means and standard deviations were generally comparable among the three groups. Over time, the means of most CBCL scales increased, while on average the standard deviations remained the same. This indicates that more problem behaviors were reported for adoptees in their later adolescent years.

The longitudinal correlations between the two assessment points show the extent to which the scores of the adoptees keep their relative positions across time, irrespective of possible changes in mean scores. For all adoptees these correlations were mostly around .60, pointing to a considerable stability of the problem behaviors over time. Only the Somatic Complaints and Thought Problems scales showed lower correlations. However, their longitudinal correlations, ranging from .32 to .64, could still be considered moderate.

The correlations between the siblings at the first assessment showed, with the exception of the Internalizing scale, that the biologically related siblings were more similar than the nonbiologically related siblings. The lower correlations for the nonbiologically related siblings suggest that genetic factors could play a role in the etiology of these behaviors. At the second assessment the correlations between the siblings, especially between the biologically related siblings, tended to be lower compared with the first assessment. For the Total Problem, Externalizing, Aggressive Behavior, Delinquent Behavior, Attention Problems and Social Problems scales the biologically related siblings still had higher correlations than the nonbiologically related siblings. This suggests that genetic factors were, also at this second assessment point, of importance. However, for all other scales the correlations of the biologically related siblings were equal to or even lower than the correlations of the nonbiologically related siblings, indicating the importance of environmental influences.

Test of ACE Model and of Means

Table IV shows the chi-squares obtained from fitting the bivariate Cholesky (ACE) model and its nested models. The first column shows the comparison between a saturated model and the full ACE model. For none of the scales did the theoretical model, specifying genetic, shared environmental and nonshared environmental factors, describe the observed data any worse than a saturated model.

In the second column, the chi-squares are given for the final ACE model. In this final ACE model, those means were constrained to be equal that did not lead to a significantly worse fit of the model. The means between the two sibling groups and the singleton group could be constrained to be equal for all CBCL syndrome scales. Between the oldest and youngest siblings the means could be constrained to be equal for all scales, with the exception of the Withdrawn scale. The oldest sibling obtained a higher score than the youngest sibling. Mean ratings of mothers and fathers at the first assessment could also be constrained to be equal for all scales, with the exception of the Somatic Complaints scale. Mothers reported more problems for

Table III. Means and Standard Deviations (between parentheses) for Oldest (First Line) and Youngest (Second Line) Siblings and Singletons, the Correlations between Siblings per Group at Assessment 1 and 2, and the Longitudinal Correlations between Assessments 1 and 2 per Group, for Achenbach's 1991 CBCL/4-18 Profiles.[a] The Number of Adoptees per Group is Given in the Last Row

	Assessment 1			Assessment 2			Longitudinal correlations		
CBCL scales	biological	nonbiological	singletons	biological	nonbiological	singletons	bio.	nonbio.	single.
Broad-band groupings:									
total problem score	11.28 (4.29)	11.79 (4.33)	11.89 (4.31)	12.56 (3.87)	12.13 (4.24)	12.38 (4.20)	.65	.66	.71
	11.49 (4.39) .57	11.56 (4.23) .39		12.26 (4.26) .45	12.03 (4.39) .31		.54	.66	
internalizing problems	7.08 (3.78)	7.04 (3.80)	6.95 (3.78)	8.39 (3.78)	7.52 (3.89)	7.78 (3.86)	.69	.59	.63
	6.17 (3.96) .31	6.43 (3.77) .34		7.41 (3.73) .17	7.28 (3.92) .40		.53	.51	
externalizing problems	6.30 (4.49)	7.03 (4.80)	7.28 (4.58)	7.29 (4.41)	7.57 (4.85)	7.71 (4.56)	.72	.71	.70
	7.43 (4.54) .47	6.92 (4.68) .19		8.00 (4.80) .40	7.26 (4.94) .18		.52	.67	
Syndrome scales:									
withdrawn	5.03 (3.22)	4.54 (3.23)	4.51 (3.25)	5.92 (3.30)	4.99 (3.35)	5.30 (3.29)	.60	.59	.61
	3.54 (3.34) .20	3.97 (3.06) .16		4.84 (3.43) .03	4.75 (3.33) .21		.46	.51	
somatic complaints	1.20 (2.21)	1.55 (2.35)	1.53 (2.30)	1.82 (2.65)	1.76 (2.55)	2.11 (2.66)	.51	.49	.38
	1.47 (2.30) .27	1.45 (2.29) .18		1.98 (2.35) .18	1.71 (2.42) .26		.37	.37	
anxious/depressed	4.24 (3.54)	4.39 (3.60)	4.26 (3.53)	5.30 (3.88)	4.80 (3.72)	4.74 (3.73)	.70	.48	.59
	3.96 (3.41) .31	3.95 (3.46) .26		4.62 (3.44) .26	4.56 (3.70) .29		.57	.51	
social problems	2.74 (2.93)	2.97 (3.15)	3.09 (3.23)	3.34 (2.93)	2.79 (3.02)	2.99 (3.13)	.57	.56	.59
	2.52 (3.06) .37	3.13 (3.29) .18		2.58 (2.95) .24	3.14 (3.34) .16		.68	.70	
thought problems	1.17 (2.06)	1.16 (2.13)	1.10 (2.03)	1.63 (2.45)	1.14 (1.98)	1.30 (2.22)	.64	.45	.38
	0.95 (2.02) .24	0.89 (1.93) .15		1.26 (2.27) .03	1.14 (2.14) .11		.41	.32	
attention problems	5.07 (3.87)	5.52 (3.70)	5.57 (3.67)	5.99 (3.19)	5.51 (3.77)	5.96 (3.60)	.65	.63	.65
	5.28 (3.73) .33	5.24 (3.92) .08		5.78 (3.47) .25	6.04 (3.77) .10		.71	.70	
delinquent behavior	2.71 (3.25)	2.85 (3.33)	3.12 (3.35)	3.69 (3.29)	3.85 (3.79)	3.84 (3.69)	.60	.63	.61
	2.97 (3.52) .43	2.67 (3.20) .24		3.92 (4.16) .33	3.34 (3.83) .17		.44	.62	
aggressive behavior	5.46 (4.08)	6.31 (4.55)	6.44 (4.33)	6.17 (4.04)	6.50 (4.53)	6.64 (4.22)	.67	.70	.69
	6.62 (4.37) .42	6.23 (4.45) .13		6.90 (4.41) .36	6.34 (4.62) .14		.51	.70	
Number of pairs/singletons	111	221	1484	76	155	1080	76	155	1080

[a] Data has been logtransformed (log 10) and multiplied by 10.

Table IV. Chi-squares Obtained from Fitting the Bivariate Cholesky ACE Model, and its Nested Models

CBCL scales	saturated vs ACE model: all means unequal df = 14	most simplified ACE model: means equal, if not leading to a significantly worse fit	ACE vs AE model df = 3	ACE vs CE model df = 3	ACE vs E model df = 6
Broad-band groupings:					
total problem score	12.634	3.672 (df = 1)	43.426[a]	9.901[a]	106.275[a]
internalizing problems	16.43	3.315 (df = 1)	39.738[a]	6.372	58.831[a]
externalizing problems	16.782	0.303 (df = 1)	11.304[a]	10.937[a]	49.501[a]
Syndrome scales:					
withdrawn	11.584	5.653 (df = 2)	11.259[a]	1.573	15.461[a]
somatic complaints	8.921	5.329 (df = 7)	15.917[a]	0.953	25.787[a]
anxious/depressed	12.586	0.386 (df = 1)	25.782[a]	7.993[a]	46.338[a]
social problems	12.931	0.001 (df = 1)	8.667[a]	5.872	27.967[a]
thought problems	19.692	0.333 (df = 1)	7.173	1.287	13.535[a]
attention problems	16.575	1.976 (df = 2)	1.689	6.869	17.463[a]
delinquent behavior	22.82	0.082 (df = 1)	15.027[a]	4.236	40.916[a]
aggressive behavior	16.63	4.683 (df = 2)	5.644	10.008[a]	33.182[a]

[a] χ^2 test of the model is significant: the model fits the observed data worse than a less constrained model. The critical χ^2 value (α = .05) with 14 df is 23.68, with 7 df is 14.06, with 6 df is 12.59, with 3 df is 7.81, with 2 df is 5.99 and with 1 df is 3.84.

this scale than fathers. At the second assessment mean ratings of mothers could not be constrained to be equal to father ratings for the Total Problem, Internalizing, Externalizing, Withdrawn, Somatic Complaints and Anxious/Depressed scales. Again, mothers reported more problems for these scales than fathers. For most scales, the means of mothers and fathers filling in one questionnaire together could be constrained to be equal to the means of other kinds of raters. Only the Somatic Complaints scale gave a significantly worse fit. Mothers and fathers filling in one questionnaire together indicated fewer problems on this scale than other raters. When still possible, the means between all four kinds of raters were constrained to be equal. For the Internalizing, Externalizing, Anxious/Depressed, Thought Problems and Delinquent Behavior scales this gave a significantly worse fit. Mothers and fathers indicated fewer problems for these scales than mothers and fathers filling in one questionnaire together or other kinds of raters. Finally, the means of the first and second assessment could only be constrained to be equal for the Social Problems scale. For all other scales the adoptees obtained significantly higher problem scores at the second assessment versus the first assessment.

The last three columns of Table IV show the chi-squares obtained from fitting the nested models of the final ACE model. To test whether the ACE model could be simplified, the loadings of the shared environmental factors were constrained at zero (ACE versus AE model). Except for the Aggressive Behavior, Attention Problems and Thought Problems scales, this model fitted the observed data significantly worse. Thus, for most scales, the shared environmental factors had to be included in the model. Second, the genetic factors were removed from the model (ACE versus CE model). Only the Total Problem, Externalizing, Aggressive Behavior and Anxious/Depressed scales showed a significantly worse fit when the CE model was compared to the full ACE model. For all other scales, the genetic factors were not statistically significant. However, one should be cautious with the interpretation of this result because this non-significantly worse fit for a model without genetic factors could have been caused by the lack of power to find genetic effects. Having almost twice as many nonbiologically related siblings as biologically related siblings, the power of this study to estimate genetic effects for the Internalizing, Withdrawn, Somatic Complaints and Thought Problems scales was low. Their power at $\propto$ = .05, assuming the estimated genetic influences at the first assessment to be true effects, was calculated to be about 25%. On the other hand, the power to detect shared

Table V. Percentage of Variance Explained by the Genetic, Shared Environmental and Nonshared Environmental Factors of the Bivariate Cholesky ACE Model and Their 95% Confidence Intervals Between Brackets

	Assessment 1	Assessment 2	Assessment 1	Assessment 2
CBCL scales	a^2	a^2	c^2	c^2
Broad-band groupings:				
total problem score	36 (5–66)	17+25=42 (1–79)	40 (28–49)	21±8=29 (15–41)
internalizing problems	16 (1–41)	8+0=8 (0–29)	30 (19–39)	29+4=33 (21–43)
externalizing problems	55 (17–90)	26+22=48 (3–88)	19 (6–30)	11+6=17 (2–29)
Syndrome scales:				
withdrawn	9 (0–41)	1+0=1 (0–29)	17 (5–27)	18+0=18 (6–28)
somatic complaints	20 (19–59)	0+0=0 (0–38)	18 (5–29)	15+9=24 (9–36)
anxious/depressed	25 (2–54)	8+0=8 (0–38)	26 (14–35)	28+0=28 (15–39)
social problems	52 (10–89)	17+0=17 (0–60)	17 (4–29)	7+6=13 (0–25)
thought problems	20 (0–22)	1+0=1 (0–33)	15 (2–25)	8+2=10 (0–23)
attention problems	53 (12–87)	19+9=28 (0–78)	7 (0–19)	5+2=7 (0–20)
delinquent behavior	34 (0–70)	34+3=37 (0–80)	25 (11–36)	8+7=15 (0–29)
aggressive behavior	61 (20–96)	37+15=52 (5–93)	13 (0–24)	7+5=12 (0–25)

Note I. Assessment 1: a^2, c^2, e^2 = percentage variance explained by first genetic, shared environmental and nonshared environmental factors at the first assessment. Assessment 2: a^2, c^2, e^2 = percentage variance explained by first + second = total genetic, shared environmental and nonshared environmental factors at the second assessment.

Note II. Some confidence intervals could not be estimated precisely because the tail of the distribution on the right side was too long.

environmental effects for these CBCL scales at $\alpha = .05$, assuming the estimated common environmental influences to be true effects, was calculated to be much higher than the power to detect genetic effects, around 89%.

As a final test of familial influences, both the shared environmental and the genetic factors were removed from the model, comparing this model with the full ACE model. For all scales this led to a significant decrease in fit, indicating that for all scales either genetic factors or shared environmental factors or both were necessary to explain the observed data.

The Estimates of the Bivariate Cholesky ACE Model

In Table V the percentage of variance explained by the genetic, shared environmental and nonshared environmental factors of the ACE model and their 95% confidence intervals are given. Estimates at the second assessment point have been divided into persistent factors that maintained their influence over time and new factors that only had an influence at the second assessment. Table VI shows the percentage of covariance between the first and second assessment, indicating what kinds of influences are responsible for the longitudinal stability of the problem behaviors. As shown in Table V, large genetic effects were found at both assessment points for the Aggressive Behavior and Externalizing scales. At the first assessment, genetic factors accounted for 61% of the variance of the Aggressive Behavior scale and 55% of the variance of the Externalizing scale. At the second assessment, genetic factors still had large effects, explaining 52% of the variance of the Aggressive Behavior scale and 48% of the Externalizing scale. Most of the genetic influences at the second assessment were caused by the continuing influences of genetic factors that had also exerted their influence at the first assessment. These persistent genetic factors maintained their importance over time, explaining at the second assessment 37% of the genetic variance of the Aggressive Behavior scale and 26% of the genetic variance of the Externalizing scale. For this last scale, new genetic factors also had a large influence at the second assessment, explaining 22% of the genetic variance. The covariance between the two assessments (Table VI) was for both scales mostly explained by genetic factors, suggesting that the stability of these scales was mostly influenced by genes, which persisted in exerting their influence over time. Shared environ-

Table VI. Percentage of Covariance between Assessment 1 and 2, Explained by the Genetic, Shared Environmental and Nonshared Environmental Factors of the Bivariate Cholesky ACE Model and Their 95% Confidence Intervals Between Brackets[a]

CBCL scales	Assessment 1–2 a^2	Assessment 1–2 c^2	Assessment 1–2 e^2
Broad-band groupings:			
total problem score	36 (−8–77)	42 (26–55)	22 (−8–59)
internalizing problems	−18 (−31–4)	48 (34–61)	70 (48–87)
externalizing problems	55 (3–101)	20 (3–35)	25 (−11–69)
Syndrome scales:			
withdrawn	−7 (−20–30)	29 (13–43)	78 (45–98)
somatic complaints	1 (−7–75)	41 (14–64)	58 (−4–104)
anxious/depressed	−24 (−40–15)	46 (30–59)	78 (43–100)
social problems	50 (−4–104)	18 (3–34)	32 (−13–80)
thought problems	−11 (−53–58)	28 (1–48)	83 (24–99)
attention problems	48 (−7–100)	9 (−4–25)	43 (−3–93)
delinquent behavior	56 (−1–107)	23 (4–40)	21 (−19–70)
aggressive behavior	69 (15–116)	14 (−2–29)	17 (−19–63)

[a] Some confidence intervals could not be estimated precisely because the tail of the distribution on the right side was too long.

mental effects were modest, explaining between 12% and 19% of the variance. At the second assessment, persistent and new shared environmental factors were almost of equal importance. Nonshared environmental factors explained between 26% and 36% of the variance. At the second assessment, new nonshared environmental factors accounted for more of the variance than the persistent factors, showing that the kinds of nonshared environmental influences that the adoptees experience probably change over time.

The Delinquent Behavior scale showed smaller genetic effects. At the first and second assessment, genetic factors explained 34% and 37% of the variance, respectively. Persistent genetic factors accounted for almost all of the genetic variance at the second assessment. The covariance was mostly explained by genetic influences indicating that also the stability of Delinquent Behavior was caused mostly by genes which maintained their influence over time. A large amount of the variance of the Delinquent Behavior scale at both assessments was also explained by nonshared environmental factors, 41% and 48% respectively. At the second assessment, these influences were mostly caused by new nonshared environmental factors, suggesting that the kinds of nonshared environmental influences on the Delinquent Behavior scale changed over time.

The Social Problems and Attention Problems scales showed large genetic effects at the first assessment, explaining 52% and 53% of the variance, respectively. However, at the second assessment the genetic factors had only moderate influences, explaining 17% and 28% of the variance, respectively. Still, almost all of the genetic influences at this second assessment were caused by persistent genetic factors. The covariance between the two assessments was for both factors mostly accounted for by genetic influences, suggesting that also the stability in having Social Problems or Attention Problems was caused mostly by genes which persisted in exerting their influence over time. Shared environmental factors were very modest for both scales, showing an almost equal influence of persistent and new factors. Nonshared environmental influences increased over the three year interval, with almost all of the variance at the second assessment accounted for by new nonshared environmental factors. This shows that also for these scales the nonshared environmental influences had changed over time.

A different pattern of results was found for the Internalizing scale, its subscales Withdrawn, Somatic Complaints and Anxious/Depressed, and for the Thought Problems scale. Although these scales showed modest genetic influences at the first assessment, almost no genetic influences were found

for any of the scales at the second assessment. The nonshared environmental factors accounted for most of the variance at both assessments. At the second assessment, new nonshared environmental factors accounted for most of the variance of the Withdrawn, Somatic Complaints and Thought Problems scales. For Internalizing and the Anxious/Depressed scale, persistent nonshared environmental influences also maintained a large influence at the second assessment. The covariance between the two assessments was for all scales mostly explained by the nonshared environmental influences, suggesting that idiosyncratic experiences were largely responsible for the stability of these scales over a 3-year interval.

The only scale having the largest percentage of covariance explained by the shared environmental factors was the Total Problem scale. The genetic, shared environmental and nonshared environmental factors accounted for 36%, 40% and 24% of the variance, respectively, at the first assessment and for 42%, 29% and 29% of the variance, respectively, at the second assessment. Almost all of the shared environmental influences at the second assessment were caused by persistent shared environmental factors. This suggests that the stability of the Total Problem scale was caused mostly by continuing influences of the same shared environmental factors.

DISCUSSION

In the present study of biologically related and unrelated adopted siblings and singletons, genetic factors are responsible for explaining a large part of the variance in the Externalizing scale at both assessment points. Persistent genetic factors, which are also expressed during the young adolescent years, maintain their importance over time, explaining 26% of the phenotypic variance at the second assessment. New genetic factors explained 22% of the phenotypic variance. The estimated covariances between the first and second assessment indicate that the genetic factors are also mostly responsible for the stability of the Externalizing scale over time. The effects of shared environmental factors are modest, showing an almost equal influence of persistent and new factors at the second assessment.

For the Internalizing scale, nonshared environmental factors account for most of the variance at both assessments. Persistent and new nonshared environmental factors account for about the same proportion of the variance during the later adolescent years. The covariance between the two assessments is mostly explained by the nonshared environmental influences, suggesting that idiosyncratic experiences are largely responsible for the stability of these scales over a three year interval. The effects of shared environmental factors are modest. At the second assessment, the persistent factors account for most of the shared environmental variance, suggesting that the familial influences for this scale do not change over time.

The longitudinal correlations, which are mostly around .60, point to a considerable stability of the problem behaviors during the three year interval. Over time, the adopted adolescents show an increase in their problem scores for all CBCL scales. This increase in problem behaviors is, according to Verhulst and Versluis-den Bieman (1995), not significantly related to either their ethnicity or to preadoption influences, like neglect, abuse, age of the child at placement in the adoptive family or medical conditions at the time of placement. Nonadopted adolescents, however, show a slight decrease in their CBCL problem scores over time. Our results indicate that the causes for stability differ for different problem behaviors. While the genetic factors are mostly responsible for the stability of the Externalizing scale, nonshared environmental factors have the largest influence on the stability of the Internalizing scale. This suggests that idiosyncratic influences, like cognitive evaluations including those related to self-esteem during adolescence, may cause the adoptees to retain high scores on the Internalizing scale. As was already concluded by Versluis-den Bieman and Verhulst (1995), adolescence, a period characterized by increasing cognitive skills, striving towards greater independence, sexual maturation and concerns over identity, may add to the problems experienced by these adopted adolescents.

Finding a low heritability for the Internalizing scale is in contrast with results obtained by twin studies (Edelbrock *et al.*, 1995; Hewitt *et al.*, 1992; Schmitz *et al.*, 1995; Van den Oord *et al.*, 1996; Van der Valk *et al.*, 1998), which show modest to large genetic effects. This difference may be due to the lack of power this study had to find genetic effects for the Internalizing scale and its subscales. We compared adopted biological and nonbiological

siblings and singletons who were all raised by their adoptive parents and who were of similar ages. This design contains more information about the shared environmental influences than about the genetic influences (Health *et al.*, 1985). The lack of power to detect genetic effects can be seen for example in the size of the confidence intervals of the estimates for genetic parameters. Estimates of genetic parameters have much larger confidence intervals than estimates of shared environmental influences, for which the power to detect effects was much larger. Another possible reason for the lack of finding genetic influences for the Internalizing scale might be that parents are less able to report on internalizing problems of adolescents. Especially with increasing age, parents are known to report fewer problems than their adolescent children do, probably because many of the internalizing problems the adolescents experience, such as anxiety and depression, remain unnoticed by their parents (Verhulst and van der Ende, 1992).

The results obtained for the Externalizing scale and its subscales are in accordance with results obtained from twin studies. The Aggressive Behavior and Externalizing scales show large genetic effects at both assessment points. Edelbrock *et al.* (1995), using the CBCL in 99 pairs of monozygotic twins and 82 pairs of dizygotic same-sex twins, ages 7–15, found that genetic factors account for 60% of the variance of the Aggressive Behavior scale and 51% of the variance of the Externalizing scale. Schmitz *et al.* (1995), also using the CBCL, found in their sample of 66 pairs of monozygotic twins and 137 pairs of dizygotic twins, mean age 8, that genetic factors explain 55% of the variance of the Aggressive Behavior scale and 57% of the Externalizing scale. Our results show that genetic factors at the first assessment, mean age 12.4 (SD of 1.2), account for 61% of the variance of the Aggressive Behavior scale and 55% of the variance of the Externalizing scale. At the second assessment, mean age 15.5 (SD of 1.2), genetic factors still have large effects, explaining 52% of the variance of the Aggressive Behavior scale and 48% of the Externalizing scale. Consistent with the other two twin studies mentioned above, shared environmental effects are modest, explaining between 12% and 19% of the variance. The genetic influences found for the Delinquent Behavior scale, the other subscale of the broad-band grouping Externalizing, are also quite similar to those obtained by Edelbrock *et al.* (1995). Their results indicate that 35% of the variance is accounted for by genetic effects, which is very close to the 34% that we have found at the first assessment. Schmitz *et al.* (1995) however, have found a much larger effect of genetic factors on the Delinquent Behavior scale, explaining 79% of the variance.

The Social Problems and Attention Problems scales both show large genetic effects at the first assessment, explaining 52% and 53% of the variance, respectively. Again these results are very similar to the results found in twin studies. In the study of Edelbrock et al. (1995), genetic factors account for 61% and 66% of the variance, respectively, while Schmitz *et al.* (1995) show influences of genetic factors of 56% and 65%, respectively. At the second assessment, however, genetic influences decrease to explaining 17% and 28% of the variance, respectively. Although this could be either a typical result of studying adopted adolescents or a normal developmental effect, it might also be that, just as with internalizing problems, parents are not well able to report on these kinds of problems for adolescents. Boomsma and Koopmans (1994) collected data on 1700 twin pairs, which were older than the twin pairs used in the studies mentioned so far (12–24 years). Using the Young Adult Self Report questionnaire (YASR; Achenbach, 1997), they found an estimated heritability of 45% for Social Problems and 42% for Attention Problems.

Only for the Total Problem scale did the shared environmental factors explain the largest percentage of the covariance between the two assessments. This is a striking result when considering that for all other scales either the genetic or the nonshared environmental factors accounted for most of the covariance. Schmitz *et al.* (1995) obtained similar results in their study, in which the Total Problem scale was one of the few scales for which the shared environmental factors explained a significant part of the variance. Van den Oord *et al.* (1996) noted that CBCL studies consistently find that shared environmental influences are larger for the Total Problem scale compared to the other scales. They hypothesized that this could be due to the fact that shared environmental influences are not expressed in a single scale but in multiple scales simultaneously.

At the second assessment, the persistent influences of the shared environmental factor explained for all problem scales the largest percentage of the

variance. Apart from the continuing influence of the familial environment, this large persistent influence could also been caused by the fact that in 62% of the sample the same kind of rater filled in the questionnaire at both assessments. Although we did allow the means of different kinds of raters and at both assessment points to differ, we could not completely correct for rater bias because we had only one completed questionnaire per child. If rater bias continued to exist in the sample, this could have enlarged the estimates of the shared environmental factor.

The CBCLs were either filled in by the mother, the father, the mother and father together or by other kinds of raters. For the Internalizing, Externalizing, Anxious/Depressed, Thought Problems and Delinquent Behavior scales, mothers or fathers filling in the questionnaire alone reported significantly fewer problem behaviors for their children than mothers and fathers together or others as raters. Although this could be a rater effect, for instance, mothers and fathers reporting more problems when they fill in a questionnaire together, it is also possible that these differences are real. Maybe parents are more concerned with the behavior of their child when the child shows more problems. Being more concerned, they probably are more likely to both take some time to answer the questionnaire. When other raters had filled in the questionnaire, most of the time the adoptee him/herself had either filled in or had assisted with filling in the questionnaire. As noted above, adolescents are known to report more problems than their parents do, especially with increasing age.

In the sample of adoptees used, 95.9% of the children were between 11 and 14 years of age at the first assessment and 95.8% of the sample was between 14 and 17 years of age at the second assessment. This does give an overlap at the age of 14. However, the overlap is small since only 18% of the sample at the first assessment and 24% of the sample at the second assessment actually constituted the group of 14-year-olds.

Because of the special sample used in this study, not all results are easily comparable with previous twin studies. In order to get sufficient power to estimate possible genetic effects on the Internalizing scale or replicate the obtained longitudinal genetic and environmental influences on the different problem behaviors, this study should be replicated with a larger, longitudinal sample of genetically informative subjects. Also, possible sex differences should be further investigated. In the adoption sample used, the group of biologically related adoptees consisted of 27% boy pairs, 31.5% girl pairs and 41.5% opposite sex pairs. The group of nonbiologically related adoptees showed a similar composition, having 20% boy pairs, 22% girl pairs and 58% opposite sex pairs. Van den Oord *et al.* (1994), using this sample at the first assessment, found small sex differences in heritability for most problem behaviors. Due to the small sample of biological siblings at the second assessment, we have not tested for sex differences in this study. However, these possible sex differences should be further examined with a larger longitudinal sample.

We estimated the likelihood based confidence intervals for all genetic, shared environmental and nonshared environmental estimates. These confidence intervals show that most genetic, shared environmental and nonshared environmental point-estimates have a rather large area in which they, depending on the precise sample and model used, can fall. Keeping this in mind, it is quite amazing that the point-estimates found in this adoption study, with the exception of the estimates of the Internalizing scale and its subscales, show so much similarity with the point-estimates found in twin studies. Both samples of genetically informative subjects have their own limitations. In twin samples, for instance, congenital anomalies are more common and parental expectations might cause the monozygotic twins to behave more alike. In adoption samples, preadoption influences and the status of "being adopted" when they grow up might cause the sample to differ from the general population. Also, in the sample of adopted children used, some background characteristics differed between the three groups. For instance, the biologically related siblings experienced more changes in their caretaking environment than the singletons, who experienced more changes again than the nonbiologically related siblings. Significant differences between the three groups were also found for age of placement in the adoptive home. Biologically related siblings were placed at an older age than singletons, who were placed again at an older age than nonbiologically related siblings. However, because the point-estimates we found show so much similarity with the point-estimates found by twin studies that do not share the same limitations as this adoption study, these results are very encouraging.

ACKNOWLEDGMENTS

This research was made possible by financial support from the Sophia Foundation for Medical Research. We thank Dr. Harold Snieder and Dr. Herma Versluis-den Bieman for their helpful advice.

REFERENCES

Achenbach, T.M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1991b). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1997). *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist*. Burlington, VT: University of Vermont, Department of Psychiatry.

Boomsma, D.I. and Koopmans, J.R. (1994). Genetic analysis of behavioral problems in young adults (abstract). *Behavior Genetics*, **24**:507.

De Groot, A., Koot, H.M. and Verhulst, F.C. (1994). Cross-cultural generalizability of the Child Behavior Checklist cross-informant syndromes. *Psychological Assessment*, **6**: (3) 225–230.

Edelbrock, C., Rende, R., Plomin, R. and Thompson, L. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *Journal of Child Psychology and Psychiatry*, **36**:(5) 775–785.

Esser, G., Schmidt, M.H. and Woerner, W. (1990). Epidemiology and course of psychiatric disorders in school-age children: results of a longitudinal study. *Journal of Child Psychology and Psychiatry*, **31**:(2) 243–263.

Graham, J.W., Hofer, S.M., Donaldson, S.I., MacKinnon, D.P., and Schafer, J.L. (1997). Analysis with missing data in prevention research. In: *The science of prevention: Methodological advances from alcohol and substance abuse research*. Bryant, K.J., Windle, M. and West, S.G. (Eds.), Washington, DC: American Psychological Association.

Graham, J.W., Hofer, S.M., and MacKinnon, D.P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, **31**: (2) 197–218.

Heath, A.C., Kendler, K.S., Eaves, L.J. and Markell, D. (1985). The resolution of cultural and biological inheritance: Informativeness of different relationships. *Behavior Genetics*, **15**:439–465.

Hewitt, J.K., Silberg, J.L., Neale, M.C., Eaves, L.J. and Erickson, M. (1992). The analysis of parental ratings of children's behavior using LISREL. *Behavior Genetics*, **22**:293–317.

Koot, H.M. (1995). Longitudinal studies of general population and community samples. In: *The epidemiology of child and adolescent psychopathology*. Verhulst, F.C. and Koot, H.M. (Eds.), London: Oxford University Press.

Little, R.J.A., and Rubin, D.B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research*, **18**:292–326.

Molenaar, P.C.M., Boomsma, D.I. and Dolan, C.V. (1991). Genetic and environmental factors in a developmental perspective. In: *Problems and methods in longitudinal research: stability and change. European Network on Longitudinal Studies on Individual Development, 5*. Magnusson, D., Bergman, L.R., Rudinger, G., Torestad, B. (Eds.), pp. 250–273. Cambridge University Press, Cambridge, England.

Muthén, B., Kaplan, D., and Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, **52**:(3) 431–462.

Neale, M.C. (1997) *Mx: Statistical modeling*. Box 126 MCV, Richmond, VA 23298: Department of Psychiatry. 4th Edition.

Neale, M.C. and Cardon, L.R. (1992). *Methodology for genetic studies of twins and families*, Kluwer Academic, Dordrecht.

Neale, M.C. and Miller, M.B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics*, **27**:(2) 113–120.

Schmitz, S., Fulker, D.W. and Mrazek, D.A. (1995). Problem behavior in early and middle childhood: an initial behavior genetic analysis. *Journal of Child Psychology and Psychiatry*, **36**:(8) 1443–1458.

Van den Oord, E.J.C.G., Boomsma, D.I. and Verhulst, F.C. (1994). A study of problem behaviors in 10- to 15-year-old biologically related and unrelated international adoptees. *Behavior Genetics*, **24**:(3) 193–205.

Van den Oord, E.J.C.G., Boomsma, D.I. and Verhulst, F.C. (1996). A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *Journal of Abnormal Psychology*, **105**:(3) 349–357.

Van der Valk, J.C.; Verhulst, F.C.; Stroet, T.M. and Boomsma, D.I. (1998). Quantitative genetic analysis of Internalizing and Externalizing Problems in a large sample of 3-year-old twins. Twin Research, 1, 25–33.

Verhulst, F.C., Akkerhuis, G.W. and Althaus, M. (1985). Mental health in Dutch children: I. A cross cultural comparison. *Acta Psychiatrica Scandinavica*, 2, Suppl. 323.

Verhulst, F.C., Althaus, M. and Versluis-den Bieman, H.J.M. (1990). Problem behavior in international adoptees: I. Epidemiological study. *Journal of the American Academy of Child and Adolescent Psychiatry*, **29**:94–103.

Verhulst, F.C., Eussen, M.L.J.M., Berden, G.F.M.G., Sanders-Woudstra, J. and van der Ende, J. (1993). Pathways of problem behaviors from childhood to adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, **32**:(2) 388–396.

Verhulst, F.C. and van der Ende, J. (1992). Agreement between parents' reports and adolescents self-reports of problem behavior. *Journal of Child Psychology and Psychiatry*, **33**:(6) 1011–1023.

Verhulst, F.C. and van der Ende, J. (1993). "Comorbidity" in an epidemiological sample: a longitudinal perspective. *Journal of Child Psychology and Psychiatry*, **34**:(5) 767–783.

Verhulst, F.C., van der Ende, J. and Koot, H.M. (1996). *Handleiding voor de CBCL/4-18*. Afdeling Kinder- en Jeugdpsychiatrie, Sophia Kinderziekenhuis, Academisch Ziekenhuis Rotterdam, Erasmus Universiteit Rotterdam, The Netherlands.

Verhulst, F.C. and Versluis-den Bieman, H.J.M. (1995). Developmental course of problem behaviors in adolescent adoptees. *Journal of the American Academy of Child and Adolescent Psychiatry*, **34**:(2) 151–159.

Versluis-den Bieman, H.J.M. and Verhulst, F.C. (1995). Self-reported and parent reported problems in adolescent international adoptees. *Journal of Child Psychology and Psychiatry*, **36**:(8) 1411–1428.

Van Westerlaak, J.M., Kropman, J.A. and Collaris, J.W.M. (1975). *Beroepenklapper*. Nijmegen, Holland: Instituut voor Toegepaste Sociologie.

Wothke, W., and Arbuckle, J.L. (1995). Full-information missing data analysis with Amos. *Paper presented at Softstat 1995*. Heidelberg, Germany, March 26–30, 1995.

Edited by Stacey Cherny